DOCUMENT RESUME

ED 312 036 PS 018 222

AUTHOR Nelsen, Mary Jo
TITLE Concurrent Validity of Three Family Day Care Assessment Instruments.
PUB DATE 16 Aug 89
NOTE 8p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Accreditation (Institutions); *Concurrent Validity; *Family Day Care; *Measurement Techniques; Program Evaluation; *Program Validation
IDENTIFIERS *Texas (San Antonio)

ABSTRACT

A research project was initiated by the San Antonio Accreditation program to compare results of overlapping validations made with three instruments: San Antonio's Family Day Care Rating Scale (FDCRS), Dallas Child Care Partnership's Family Day Care Observation Instrument (FDCOI), and the National Association for Family Day Care's (NAFDC) Assessment Profile for Family Day Care. During April-August, 1989, a total of 32 validations were conducted for the San Antonio Accreditation program by two validators. Validators evaluated each provider or home concurrently. One observer used the FDCRS and the FDCOI to rate providers and homes, while the second observer used the FDCRS and the NAFDC. Validators were professionals in the field of early childhood education. Concurrent validity was not established for the NAFDC. Of 32 providers, 4 failed their first attempt at accreditation according to city standards. Of these 4, two would have also failed the Dallas and NAFDC requirements for accreditation; the third would have failed only the Dallas instrument; and the fourth would have passed the NAFDC and Dallas scales. Research is needed to establish the validity of the NAFDC scale. (RH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED312036

CONCURRENT VALIDITY OF THREE FAMILY DAY CARE ASSESSMENT INSTRUMENTS

BY MARY JO NELSEN
8/16/89

PS 018222

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Mary Jo Nelsen

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

CONCURRENT VALIDITY OF THREE FAMILY DAY CARE ASSESSMENT INSTRUMENTS

Currently, three independent accreditation programs operate to recognize quality family day care providers. The original family day care accreditation program, established in 1986, is operated by the City of San Antonio's Childrens Resources Division and uses the Family Day Care Rating Scale (FDCRS) (Harms and Clifford, 1985) as the quality assessment instrument. In 1988 Dallas Child Care Partnership and the National Association for Family Day Care (NAFDC) initiated accreditation programs which used newly developed assessment instruments. Accurate assessments of quality are crucial in a responsible family day care accreditation program. Parents base decisions regarding child care on many criteria, however, when providers say they are accredited- it must mean quality care will be provided for children. Otherwise, children are left vulnerable to inappropriate care and possible harm while accreditation systems risk legal culpability. Therefore, the validity of these assessment instruments is of real concern to the field of early childhood education and practitioners attempting to improve quality in family day care homes through accreditation. For this reason, a research project was initiated by the San Antonio Accreditation program to compare results of overlapping validations using the three instruments currently available.

The FDCRS has been used for three years in the San Antonio Accrediation Program with over 200 validations having been conducted- generating considerable data on provider performance prior to and after training. Results of an item analysis of sixty-two validations were released in a recent report (Nelsen, 1989) which showed significant increases between pretest and posttest scores on the FDCRS. Additionally, the FDCRS has been used in numerous research projects as a measure of family day home quality (Altman and Mills, 1988; Fischer, 1989; Goelman and Pence, 1987; Howes and Stewart, 1987; Jones and Meisels, 1987; Pence and Goelman, 1987; Pepper and Stuart, 1985; and Stuart and Pepper, 1985). The FDCRS was developed in 1984 at the University of North Carolina as an adaptation of the Early Childhood Environment Rating Scale (ECERS) (Harms and Clifford, 1980). Content validity of the FDCRS is based on that of the ECERS. Content of the FDCRS has been correlated with that of the Child Development Associate (CDA) Family Day Care credential, which became available nationally in 1985. With proper observer training uses of the FDCRS consistently achieve a satisfactory level of interrater reliability (85-95%). Internal consistency for the six subscales was calculated using Cronbach's alpha and ranged from .70 to .90 (Harms and Clifford, 1988).

The FDCRS is a Likert scale composed of 33 items over six categories (subscales); Space and Furnishings, Basic Care, Language and Reasoning, Learning Activities, Social Development, and Adult Needs. Each item is described on four levels of quality: inadequate (does not meet even custodial needs), minimal (meets custodial needs and to some extent basic developmental needs), good (meets developmental needs), and excellent (high quality personalized care). Scores of 1-7 are assigned for each of the 33 items with a score of "1" corresponding to inadequate care

and "7" to excellent care. Intermediate scores of "2","4", and "6" are awarded when the provider meets all of the requirements of the lower score and some of the requirements of the higher score. A slightly revised version of the FDCRS has recently been published (Harms and Clifford, 1989). Recommended length of observations for this instrument is two hours.

The instrument used in the Dallas accreditation program is called the Family Day Care Observation Instrument. This instrument was designed around CDA criteria to assure criterion validity. Interrater reliability was established for the Dallas instrument (97.5%) during the development phase by comparing apparently non-concurent ratings on nine homes by two observers. Content validity for the Dallas instrument was established by referencing the literature and through an expert panel convened during the field test. Also, a weighting system for the subscales was incorporated into the accreditation decision process. Providers were expected to score 90% or greater on the Health, Safety and Nutrition section, an 80% or greater on four sections combined which cover Activities and 50% or greater on three sections; Business, Parents, and Environment. This weighting system effectively converts the eight sections of the Dallas instrument to five by combining the last four sections on Activities (for physical, self-social, language-thinking, and creative development). The instrument is composed of 109 items which the observer rates by recording either a "yes", "no", or "not applicable" response. On several of the items a list of criteria are provided and the observer must find all or a specified number of criteria are met to mark a "yes". In a few items examples of ways of meeting the criteria are given and the observer must decide if the provider qualifies for a "yes". Recommended length of each observation is four to six hours.

FIGURE 1: Subscales/Dimensions For Each Scale and Numbers of Items

FDCRS	NAFDC	DALLAS
(6) Space and Furnishings	(34) Indoor Safety	(9) Business Management
(7) Basic Care	(27) Health	(6) Parents/ Families
(4) Language & Reasoning	(16) Nutrition	(13) Learning Environment
(9) Learning Activities	(22) Indoor Play Environment	(43) Safety, Health, Nutrition
(4) Social Dev.	(49) Interacting	(38) Physical, Social, Lang/Think., Creative
(3) Adult Needs	(14) Outdoor Play Environment	
	(27) Professional Responsibility	
(33) TOTAL	(189) TOTAL	(109) TOTAL

The assessment instrument used in the NAFDC accreditaion program is the Assessment Profile for Family Day Care (Sibley and Abbott-Shim, 1987). The instrument and accreditation system were field tested in seven states with 28 providers. Information on validity and reliability is not currently available on this scale, although the authors state that they are monitoring interrater reliability between the three observers (parent, provider and validator) as the instrument is used in accreditations and are finding over 97% agreement. The NAFDC scale consists of 189 items in seven dimensions: Indoor Safety, Health, Nutrition, Indoor Play Environment, Interacting, Outdoor Play Environment and Professional Development. Items are scored "yes" or "no" with any "not applicable" responses counted as a "yes". Length of observations are supposed to be six hours. Although the actual accreditation decision is based on many additional factors, the minimum acceptable score on the instrument is no less than a 75% on any category and no less than an 85% composite score. The scale itself accounts for 80% of the accreditation decision.

Although, the two newly developed assessment instuments look very different from the FDCRS due to the surface difference between a Likert scale and a "yes"/"no" format, the results of overlapping assessments should be consistent, since, the purpose of the three scales is the same- namely, to identify quality family day care. If concurrent validity of the Dallas and NAFDC instruments can be established by comparison to the FDCRS in this study, then it may be said that the three family day care assessment instruments measure the construct of quality child care. Although, subscale comparisons between the instruments would be interesting they were not attempted since each instrument categorizes items differently. For example, the FDCRS has no item for interaction between caregiver and children but these interactions are assessed as a part of numerous items throughout the scale, while, the NAFDC instrument has a separate dimension (or subscale) for interactions (See Fig. 1). Consequently, this study examines total scores on the three assessment instruments. This study was not designed to compare the three accreditation systems.

METHOD

During April through August of 1989 thirty-two validations for the San Antonio Accreditation program were conducted where two validators evaluated each provider/home concurrently. One observer rated the provider/home using the FDCRS and the Dallas instrument, while, the second observer used the FDCRS and the NAFDC instrument. Validator/observers were all professionals in the field of early childhood education (a center director, a local training consultant, a CDA and accredited day home provider, and a day care licensing expert) as well as two early childhood graduate students from a local college. Six validators had been selected from a training group of fifteen individuals who received classroom training as well as in-home practice using the scales. Observation techniques and scoring methods which the validators were instructed to use with each of the scales were based on suggestions from each of the scale's authors. Interrater reliability during validator training measured 89-95%. Because of overlapping of the FDCRS assessments we were able to monitor interrater reliability

throughout the study. We operated under the assumption that if adequate interrater reliability was maintained on the FDCRS that the validator's agreement on what they were observing would be good and their scoring of the NAFDC and Dallas instruments would be based on the same evaluations of the same events. During the study interrater reliability was 85%, which is considered acceptable for the FDCRS. For both the NAFDC and Dallas instruments scores were expressed in percentages since total possible scores varied. The scores that were used were the mean of each subscale score. This was done in order to equalize any possible differences between assessment instruments in the natural weighting of subscales by virtue of the differing numbers of items within each area. Pearson's r correlations were conducted to compare the total scores for the three scales and to determine concurrent validity. Although coeffecients at or above .90 would be excellent, the level selected for acceptablility in this study was set at .80 or greater. Mean total scores and standard deviations were also calculated.

RESULTS

All correlations achieved extremely high levels of significance (alpha <.001) showing that the relationships found are true and not random. However, correlation coeffecients showed mixed results on validity for the instruments (see Chart A). The validity coeffecient found for the Dallas instrument was .82, an acceptable level to indicate validity. The Dallas instrument appears to measure quality child care as well as the FDCRS does. However, the NAFDC coeffecient measured an unacceptably low .62 and fails to establish concurrent validity for the instrument. When the Dallas and NAFDC instruments were compared with Pearson's r the resulting coeffecient was .62.

CHART A: Correlational Coeffecients with Pearson's r

Scales compared	Coeffecient	Significance
Dallas/FDCRS	.82	<.001
NAFDC/FDCRS	.62	<.001
Dallas/NAFDC	.62	<.001

Mean scores and standard deviations group were calculated to look for any indications of possible simplified scoring techniques for future revisons to the San Antonio Accreditation system. The mean score on the FDCRS was a 5.4 (S.D.=.7), on the Dallas instrument the mean was 87.2% (S.D.=6.9%), and on the NAFDC scale the mean score was 92.5% (S.D.=4.3%) (see chart B).

CHART B:	Mean Score	Standard Deviation
FDCRS	5.40	.72
Dallas	87.19%	6.94%
NAFDC	92.47%	4.28%

DISCUSSION

This study was not designed to evaluate the differences in the three accreditation systems which utilize the three assessment instruments, however, some comparisons between hypothetical and real pass/fail distinctions for particular providers included in the study are interesting. Four providers (out of 32) failed their first attempt at accreditation according to City standards which require all "3's" or higher except for three items which may receive "2's". Of these four providers; two would have also failed the Dallas and NAFDC requirements for accreditation, one would have also failed only the Dallas instrument and the fourth provider would have passed the NAFDC and Dallas scales. The provider who would have only failed the FDCRS points out one real difference between the practical effects of a Likert scale and a "yes"/"no" scale. This provider earned a "1" for the item "Use of Television" by keeping her TV on throughout the observation causing an automatic fail. Neither of the newer scales reacts so harshly to this type of event. Any interpretation of these results warrant caution, however, since NAFDC and Dallas accreditation decisions are based on more than the minimum scores quoted earlier in this paper and applied here to compare results for these four providers. However, it is interesting to note that again the FDCRS seems to be mirrored more closely in the Dallas instrument than the NAFDC scale.

Much of the inconsistency in results on the three scales can be attributed to differences in items. Neither NAFDC or Dallas instrument has items on multicultural materials/activities, while one of the areas that providers frequently score low on with the FDCRS is this area. Under the 1985 FDCRS providers without CPR could score a "7" on items for Health and Safety (revisions to the scale require CPR training for a "7").

It is difficult to explain the failure of this study to establish concurrent validity for the NAFDC instrument. One possible explanation is that the study was somehow flawed, however, this would be more likely if none of the results were positive. Future research is needed to establish validity for the NAFDC scale, perhaps, as the NAFDC accreditation system develops this can be done. Although the Dallas instrument is the least polished in appearance of the two newer scales, it apparently does an acceptable job of measuring family day care quality. Neither of the newer scales takes less time to administer than the FDCRS, however, organizations have chosen to develop and utilize the longer scales and the first question is whether or not they are valid. The improvement of and support for family day quality in this country is an important task, which accreditation systems may contribute toward. This, however will require the existence of a solid research base in support of the development of family day care assessment instruments used in these accreditation systems.

BIBLIOGRAPHY

Altman, J. S. and Mills, B. C. (1988). Caregiver behaviors and adaptive behavior development of very young children in home care and daycare. ED291513. ERIC.

Child Care Partnership of Dallas (1987). Family Day Home Observation Instrument. Dallas, Texas.

Fischer, J. L. (1989). Family day care: factors influencing the quality of caregiving practices. unpublished doctoral dissertation.

Goelman, H. and Pence, A. R. (1987). Effects of child care, family and individual characteristics on children's language development: the Victoria day care research project. in Quality in Child Care: What Does Research Tell Us? ed. by Deborah A. Phillips. pp89-104.

Harms, T. and Clifford, R.M. (1980). Early Childhood Environment Rating Scale. Teachers College Press, Columbia University, New York.

Harms, T. and Clifford, R.M. (1985). Family Day Care Rating Scale. University of North Carolina.

Harms, T. and Clifford, R.M. (1988). The Family Day Care Rating Scale. unpublished paper presented at the 1988 annual conference of the National Association for the Education of Young Children.

Harms, T. and Clifford, R.M. (1989). Family Day Care Rating Scale. Teachers College Press, Columbia University, New York.

Howes, C. and Stewart, P. (1987). Child's play with adults, toys and peers: an examination of family and child care influences. Developmental Psychology. 23 (3), 423-430.

Jones, S. N. and Meisels, S. J. (1987). Training family day care providers to work with special needs children. Topics in Early Childhood Special Education, 1 (1), 1-12.

Nelsen, M.J. (1989). Does training make a difference? Unpublished paper. submitted to ERIC.

Pence, A. R. and Goelman, H. (1987). Who cares for the child in day care? An examination of caregivers from three types of care. Early Childhood Research Quarterly. 2 (4), 315-334.

Pepper, S. and Stuart, B. (1985). Informal family day care: a study of caregivers. Ontario Mental Health Foundation.

Sibley, A. and Abbott-Shim, M. (1987). Assessment profile for family day care. Quality Assist; Georgia.

Stuart, B. and Pepper, S. (1985). Private home day care providers in Ontario: a study of their personal and psychological characteristics. Ontario Mental Health Foundation.